DOCUMENT RESUME

ED 128 340 95 SP 010 412

AUTHOR Stallings, Jane
TITLE A Study of Implementation in Seven Follow Through Educational Models and How Instructional Processes Relate to Child Outcomes.
INSTITUTION Stanford Research Inst., Menlo Park, Calif.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Nov 75
NOTE 35p.; Paper presented at the National Invitational Conference on Research on Teacher Effects: An Examination by Policy-Makers and Researchers (Austin, Texas, November 3-5, 1975)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Disadvantaged Youth; Educational Innovation; Educational Research; Experimental Curriculum; Experimental Schools; *Experimental Teaching; *Instructional Programs; *Primary Education; *Program Development; *Program Evaluation; Teaching Methods
IDENTIFIERS *Project Follow Through

ABSTRACT

The purpose of the Follow Through Classroom Observation Evaluation was to assess the implementation of seven Follow Through sponsor models included in the study and to examine the relationships between classroom instructional processes and child outcomes. The seven programs selected for study include two behavioristic models, an open school model based upon English Infant School Theory, and three other models based on particular combinations of theory and practice drawn from Piaget, Dewey, and the English Infant Schools. To study sponsor implementation, two questions were asked: (1) are the individual models consistently implemented in accordance with the sponsor prestated philosophies and objectives? and (2) do meaningful differences as planned exist among the individual sponsor models; that is, have the planned variations actually been achieved? Another question asked is central to the primary objectives of the follow through evaluation: How are children affected by the different approaches within these planned educational programs? Four first- and third-grade classrooms were observed in 36 towns and cities. The projects included in the sample represented all geographic regions, urban and rural areas, and several racial and ethnic groups. The SRI Classroom Observation Instrument was employed to gather data about classroom environment and processes. It consists of five sections: (1) Classroom Summary Information; (2) Physical Environment Information; (3) Classroom Checklist; (4) Preamble, and (5) Five Minute Observation. The study concludes that what occurs within a classroom does contribute to achievement in basic skills, good attendance, and desired child behaviors. (DMT)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED128340

SRI International
Menlo Park, California 94025 · U.S.A.

A STUDY OF IMPLEMENTATION IN SEVEN FOLLOW THROUGH EDUCATIONAL MODELS AND HOW INSTRUCTIONAL PROCESSES RELATE TO CHILD OUTCOMES

Jane Stallings, Ph.D.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Presented at the Conference on
Research on Teacher Effects:
An Examination by Policy-Makers and Researchers

Sponsored by the National Institute of Education

University of Texas at Austin
November 3-5, 1975

SP010 412

MAJOR FINDINGS*

- Seven sponsors' educational models were studied in five locations each. Each model had unique features that were consistently present in all five of a sponsor's sites. The seven sponsors' models are considered to be well implemented and could be expected to be implemented in other locations similar to those in the study.

- Classroom instructional practices in first and third grade classrooms have been identified that contribute to higher test scores, other desired child behaviors, and lower absence rates.

- Highly controlled classroom environments in which teachers used systematic instruction and a high rate of positive reinforcement contributed to higher scores in math and reading. Flexible classroom environments that provide more exploratory materials and allow for more choice on the part of the child contributed to higher scores on a test of nonverbal reasoning, lower absence rates, and a willingness on the part of children to work indeperdently.

---

* These findings are supported by statistical analysis in Follow Through Classroom Observation Evaluation 1972-1973, Jane Stallings and David Kaskowitz, Stanford Research Institute, Menlo Park, California, August 1974.

A STUDY OF IMPLEMENTATION IN SEVEN FOLLOW THROUGH EDUCATIONAL MODELS AND HOW INSTRUCTIONAL PROCESSES RELATE TO CHILD OUTCOMES

Introduction

The purpose of the Follow Through classroom observation evaluation was to assess the implementation of the seven Follow Through sponsor models included in the study and to examine the relationships between classroom instructional processes and child outcomes. An important question continually raised by government agencies and educators has been "Does the Follow Through Program of Planned Variations exist?" (i.e., are a variety of educational programs being successfully implemented and do these models differ significantly from each other?) The Follow Through Program was established by Congress in 1967 under the Office of Economic Opportunity when it became apparent that a program was needed in the early grades of public school to reinforce and extend the academic gains made by economically disadvantaged children enrolled in Head Start or similar preschool programs. Project Follow Through was and is a "planned variation" research design; that is, the goal was to examine the differential effectiveness of programs based on divergent educational and developmental theories. The program began when researchers and other educational stakeholders were invited by the government to submit plans for establishing their various programs in public schools in order to test whether their individual approaches could improve the educational achievement of economically disadvantaged children. The seven programs selected for study in this analysis represent a wide spectrum of innovative educational theories represented in Follow Through. The range includes two more behavioristic models (the University of Kansas and the University of Oregon), a model based upon the theory of Piaget (High/Scope), an open school model based upon English Infant School Theory (EDC), and three other models which each have their own particular combinations of theory and practice drawn from Piaget, Dewey, and the English Infant Schools (Far West Laboratory, University of Arizona, and Bank Street).

To study sponsor implementation, two questions were asked: (1) are the individual models consistently implemented in accordance with the sponsor prestated philosophies and objectives? and (2) do meaningful differences as planned exist among the individual sponsor models; that is, have the planned variations actually been achieved? The third question asked is central to the primary objective of the Follow Through evaluation: "How are children affected by the different approaches embodied within these planned educational programs?"

The answers to these questions are not only basic to the overall evaluation of Follow Through, but are also of importance to future planning of educational intervention programs. An affirmative answer to the first question would help establish that educational innovations have been introduced as planned and that the Follow Through program has met its goal of planned variation. An affirmative answer to the question of implementation would help establish that Follow Through models are exportable to other locations and that the sponsors have established meaningful delivery system that enable them to implement their models consistently in diverse locations. Finally, information regarding how classroom instructional processes used in the various programs impact upon the growth and development of children is of critical importance to all educators who plan school programs.

The present summary represents highlights of the full report prepared by Stanford Research Institute, Menlo Park, California--Follow Through Classroom Observation Evaluation 1972-1973.[1]

Sample

Four first grade and four third grade classrooms were observed in 36 cities and towns. This represented five projects for six Follow Through educational models and six projects for University of Arizona's model. One first grade and one third grade Non-Follow Through classroom were selected for comparisons at each project. These Non-Follow Through classrooms were combined to form a pooled comparison group. The projects included in the sample represented all geographic regions, urban and rural areas, and several racial and ethnic groups (see Table 1).

Measurements

Classroom Processes

The SRI Classroom Observation Instrument was employed to gather data about classroom environment and processes. The instrument was initially developed in 1969 with the assistance of eight Follow Through sponsor representatives with a goal of being flexible enough to record the salient features of a variety of program components.

The instrument consists of five sections:

- Classroom Summary Information (CSI)--The CSI is filled out once each day. It identifies the sponsor and teacher and provides information on the number of teachers, aides, volunteers, and students, and the class duration.

Table 1

CLASSROOM OBSERVATION SAMPLE, SPRING 1973

Sponsor and Sites		Number of Follow Through Classes Observed: First Grade	Third Grade
Far West Laboratory for Educational R&D			
0201	Berkeley, Calif.*	4	4
0204	Duluth, Minn.*	4	4
0207	Lebanon, N.H.	4	4
0209	Salt Lake City, Utah	4	4
0213	Tacoma, Wash.	4	4
University of Arizona			
0305	Des Moines, Iowa	4	4
0307	Fort Worth, Texas*	4	4
0308	LaFayette, Ga.	3	4
0309	Lakewood, N.J.	4	4
0311	Newark, N.J.	4	4
0316	Lincoln, Nebraska	4	4
Bank Street College			
0502	Brattleboro, Vermont	3	3
0504	Fall River, Mass.	4	4
0506	New York City, P.S. 243K	4	4
0508	Philadelphia II, Pa.*	4	4
0510	Tuskegee, Ala.*	4	4
University of Oregon			
0703	E. St. Louis, Ill.	4	4
0707	New York City, P.S. 137K	3	3
0708	Racine, Wisc.	4	4
0711	Tupelo, Miss.*	4	4
0719	Providence, R.I.	4	4
University of Kansas			
0801	New York City, P.S. 77X*	2	2
0803	Philadelphia VI, Pa.*	4	4
0804	Portageville, Mo.*	4	3
0806	Kansas City, Mo.	4	4
0807	Louisville, Ky.	4	4
High Scope Educational Research Foundation			
0901	Greenwood, Miss.*	4	4
0902	Ft. Walton Beach, Fla.*	4	4
0903	New York City, P.S. 92M	3	4
0906	Greeley, Colo.	3	3
0907	Denver, Colo.	4	4
Education Development Center			
1101	Burlington, Vermont	4	4
1103	Philadelphia IV, Pa.*	4	4
1106	Paterson, N.J.*	4	4
1107	Rosebud, Texas	3	3
1108	Smithfield, N.C.	4	2
	Total	136	135

*These sites have been observed previously.

- Physical Environment Information (PEI)--The PEI is filled out once each day. It provides information on the seating patterns and on the presence and use of equipment and materials.

- Classroom Check List (CCL)--A CCL is filled out about four times an hour. It provides information on the grouping of children and teaching staff and activities in the classroom.

- Preamble (PRE)--A Preamble is filled out subsequent to each CCL. It contains information about the activity and role of the person who is the focus of the FMO interactions.

- Five-Minute Observation (FMO)--The FMO is filled out subsequent to each Preamble. It contains information in the form of coded sentences concerning the type of interactions occurring in the classroom. The information includes the parties to the interaction, the type of interaction, and the quality of the interaction.

Four dimensions of reliability have been examined in the main report of Stallings and Kaskowitz: day-to-day stability of classroom processes, observer reliability, confusability of the operational definitions of the observation codes, and anomalies in the data collected. Classrooms were found to be acceptably stable on observed variables from one day to another. Codes found to be unreliable were omitted from further study.[2] Anomalies in the data were deleted where warranted; for example, if the teacher went home sick in the middle of the morning, that day's observation was deleted.

Child Measures

The entering ability of the children was assessed by the Wide Range Achievement Test (WRAT) which was administered to the children when they first entered school, either at the kindergarten or first-grade level.

Reading and math skills were assessed by the Metropolitan Achievement Test (MAT) in both first and third grades.

Problem-solving skills (perceptual) were assessed in third grade only, using the Raven's Coloured Progressive Matrices (Raven's). This test was designed by John C. Raven (1956) as a culture-fair test of nonverbal reasoning, or fluid problem-solving ability in visual perceptual tasks.

The Intellectual Achievement Responsibility Scale (IAR), used in the third grade only, assessed the extent to which the child takes responsibility for his own successes or failures or attributes his achievements to the operation of external forces (e.g., luck or fate).

Child behaviors were assessed through systematic observations recorded on the SRI Observation Instrument. The behaviors reported here are independence, task persistence, cooperation, and question asking.

Absences from school were determined from school records.

Methodology

The first step in the assessment of classroom implementation was to describe each educational model in detail. The model descriptions were prepared by SRI and reviewed by the sponsors and then revised according to the sponsor's specifications. The second step was to create variables from the codes used in the observation instrument which would describe representative elements of each sponsor's model. Each sponsor identified these variables which were (1) important to his model and (2) expected to occur more frequently than in conventional classrooms. A list of variables was selected for each of the seven models. These ranged in number from 16 for University of Oregon to 28 for Far West Laboratory (see Table 2). It is noted that the critical list of variables describes a sponsor's model only in part. The observation instrument employed in the study is not designed to capture the important subtle processes of some of the programs as, for example, that of developing intrinsic motivation.

Since the Follow Through programs are intended to be innovative and to represent alternatives to the conventional classroom, a pool of Non-Follow Through classrooms was used as the standard from which Follow Through classrooms were expected to differ in specified ways. The standards were established separately for first and third grades.

In observational data the distribution of scores rarely follows a normal curve; thus, a non-parametric scaling techinque was used in the implementation analysis. Implementation scores for each sponsor were determined by rank ordering the Non-Follow Through classroom mean scores on each sponsor variable and dividing the distribution into five equal parts or quintiles. Each sponsor classroom has a score on each variable and falls within a quintile range. A sponsor's implementation score on any variable will always be a score between 1 and 5. This represents the position of a Follow Through classroom mean relative to the distribution of Non-Follow Through means. A total implementation score was computed for each classroom, each project, and each sponsor.

In order to assess the magnitude of the total implementation scores for Follow Through classrooms, a total implementation score was also computed for each Non-Follow Through classroom on each sponsor's set of implementation variables. The mean and standard deviation of the Non-Follow Through pooled classrooms are reported for each sponsor separately for first and third grades. One-tailed t tests were computed to test for the significance of the differences between each Follow Through sponsor's classrooms and the Non-Follow Through classrooms. Analyses of variance were also computed to examine the within-site and among-site difference in total implementation scores for each sponsor. Implementation was judged on two criteria: (1) do the sponsored classrooms differ significantly from Non-Follow Through? and (2) are the classrooms similar in implementation both within projects and among projects?

Table 2

LIST OF CRITICAL VARIABLES SELECTED BY SPONSORS

No.	Variables Description	Far West Labs	University of Arizona	Bank Street	University of Oregon	University of Kansas	High Scope	EDC
24	Child selection of seating and work groups	X	X	X			X	X
25	Games, toys, play equipment present	X	X	X			X	X
30	General equipment, materials present	X				X		X
65	Guessing games, table games, puzzles		X			X	X	
66	Numbers, math, arithmetic	X	X	X	X	X	X	X*
67	Reading, alphabet, language development	X	X	X	X	X	X	X*
70	Sewing, cooking, pounding		X				X	
71	Blocks, trucks						X	
74	Practical skills acquisition	X**		X				
81	Wide variety of activities, over one day	X	X	X			X	X
86	Teacher with one child	X	X	X			X	
87	Teacher with two children			X				
88	Teacher with small group		X	X	X	X	X	
92	Aide with one child	X		X				
94	Aide with small group		X		X	X	X	
114	One child independent	X	X	X				X
115	Two children independent							X
116	Small group of children independent	X		X				X
139	Math or science equipment/Academic Activities	X		X			X	X
140	Texts, workbooks/Academic Activities				X	X		
141	Child to adult, all verbal except response			X				
150	Individual child verbal interactions with adult	X	X	X	X	X	X	X
160	Child questions to adults	X	X	X			X	X
161	Child group response to adult academic commands/requests or direct questions				X			
172	Child presenting information to a group			X			X	
175	Adult instructs an individual child	X						X
176	Adult instructs a group				X			
190	Adult task-related comments to children						X	X
196	All adult acknowledgment to children	X		X	X	X	X	
198	All adult praise to children		X		X	X		
412	Adult feedback to child response to adult academic commands/requests, questions				X	X		
420	Adults attentive to a small group	X			X		X	
421	Adults attentive to individual children	X	X	X		X	X	
423	Positive behavior, adults to children	X	X	X				
435	Total academic verbal interactions				X			
438	Adult communication or attention focus, one child	X		X		X	X	X
440	Adult communication or attention focus, small group				X		X	
446	Adult movement	X						
450	All child open-ended questions							X
451	Adult academic commands/requests and direct questions to children				X	X		
452	Adult open-ended questions to children	X	X	X			X	X
453	Adult response to child's question with a question						X	X
454	Child's extended response to questions		X	X				
456	All child task-related comments	X	X				X	
457	All adult positive corrective feedback	X			X	X		
460	All child positive affect	X	X					X
469	All adult reinforcement with tokens					X		
509	Child self-instruction, academic				X*			
510	Child self-instruction, objects			X			X	X
511	Child task persistence			X		X	X	
514	Two children working together, using concrete objects						X	
515	Small group working together, using concrete objects			X			X	
516	Social interaction among children	X		X				X
574	Child movement	X						X
599	Child self-instruction, nonacademic	X	X				X	
	Total number of Critical Variables	28** 27*	24	27	16** 17*	17	29	20** 22*

* Third grade only.

** First grade only.

Results

As shown in Table 3, both first and third grades of the Responsive Educational Program of Far West Laboratory are significantly different in implementation from Non-Follow Through first and third grades. The projects assessed were Berkeley, California; Duluth, Minnesota; Lebanon, New Hampshire; Salt Lake City, Utah; and Tacoma, Washington. Far West Laboratory's classrooms are remarkably similar both within sites and between sites. The greatest difference was found in the third grades of Duluth.

University of Arizona classrooms differ significantly from Non-Follow Through classrooms in both first and third grades. However, there is quite a variation among the mean scores of projects. The projects assessed were Des Moines, Iowa; Fort Worth, Texas; LaFayette, Georgia; Lakewood, New Jersey; Newark, New Jersey; and Lincoln, Nebraska. The highest first grade scores were found in Lincoln (89.5) and the lowest in Newark (54.0). Although there is little deviation among classrooms within the latter site, both first grade and third grade implementation scores are lower than the Non-Follow Through scores. The sponsor noted some staffing problems which may have limited the implementation level at Newark. It must be noted that the other Arizona sites had implementation scores at least ten points higher than the Non-Follow Through scores (see Table 4).

The Bank Street first and third grade classrooms differ significantly from Non-Follow Through in implementation scores. Further, there is little deviation either within projects or among projects at either grade level. The projects assessed were Brattleboro, Vermont; Fall River, Massachusetts; New York City, New York; Philadelphia, Pennsylvania; and Tuskegee, Alabama. The greatest deviation is found among third grade classrooms in Tuskegee (mean 73.0, S.D. 8.1) and among Fall River first grade classes (mean 72.6, S.D. 7.5). The classrooms in Tuskegee are scattered over a large county area and consistency among classroom processes may be hard to obtain. In Fall River, which has a large Portuguese population, 34 percent of the children do not speak English as a first language. Some of the classrooms seem to have many Portuguese-speaking children while other classrooms have few such children, and such differences may affect implementation scores if the teacher has difficulty in understanding the children or making herself understood. Whether or not the classroom with a low implementation score also has more children who do not use English as a first language has not been investigated. Even though the deviation between classrooms is greater, both Tuskegee and Fall River have implementation scores considerably higher than Non-Follow Through. Only Philadelphia third grade classrooms have a mean score close to that of the Non-Follow Through classrooms. 1972-73 was a difficult year for sponsors to work in Philadelphia due to two major teacher strikes (see Table 5).

Overall the University of Oregon's classrooms are significantly different from the Non-Follow Through classrooms. The F test used in the analysis of variance indicated that the among-site variance is greater than the within-site variance in the first grade. The projects assessed were E. St. Louis, Missouri; New York City, New York; Racine, Wisconsin; Tupelo,

Table 3

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--FAR WEST LABS

First Grade

Sites		Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Berkeley	(EK)	72.6%	79.3%	75.6%	71.9%	74.8%	3.4
Duluth	(EK)	76.3	84.4	80.7	80.0	80.4	3.3
Lebanon	(EK)	81.5	75.6	84.4	80.7	80.6	3.7
Salt Lake City	(EK)	80.7	85.9	75.6	80.0	80.6	4.2
Tacoma	(EK)	78.5	71.9	78.5	71.1	75.0	4.1
Sponsor Scores (N = 20):						78.3%	4.4
NFT Scores (N=35):						60.3	6.3
						t = 11.28	
						p < .001	
						f = 2.65	
						p < NS	

Third Grade

Sites		Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Berkeley	(EK)	82.2%	70.4%	79.3%	71.9%	75.9%	5.7
Duluth	(EK)	74.1	61.5	80.0	71.1	71.7	7.7
Lebanon	(EK)	69.6	77.8	74.1	64.4	71.5	5.8
Salt Lake City	(EK)	84.4	89.6	76.3	85.2	83.9	5.6
Tacoma	(EK)	79.2	84.4	80.7	72.6	79.2	5.0
Sponsor Scores (N=20):						76.4%	7.2
NFT Scores (N=36):						59.0	9.4
						t = 7.18	
						p < .001	
						f = 3.07	
						p < .05	

(NFT means Non-Follow Through)

Table 4

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--UNIVERSITY OF ARIZONA

	First Grade					
	Classroom Scores				Site Scores	
Sites	1	2	3	4	$\bar{X}$	S.D.
Des Moines (EK)	79.0%	62.9%	69.5%	71.4%	70.7%	6.7
Fort Worth (El)	85.7	78.1	70.5	75.2	77.4	6.4
LaFayette (El)	79.0	71.4	87.6		79.4	8.1
Lakewood (EK)	78.1	74.3	76.2	79.0	76.9	2.1
Newark (EK)	57.1	54.0	56.2	54.3	55.4	1.5
Lincoln (EK)	89.5	88.6	74.3	81.0	83.3	7.1
Sponsor Scores (N=23):					73.6%	10.7
NFT Scores (N=35):					61.8	7.0
					t = 4.99	
					p < .001	
					f = 11.76	
					p < .001	

	Third Grade					
	Classroom Scores				Site Scores	
Sites	1	2	3	4	$\bar{X}$	S.D.
Des Moines (EK)	65.7%	52.4%	53.3%	75.2%	61.7%	10.9
Fort Worth (El)	66.7	80.0	82.9	84.8	78.6	8.2
LaFayette (El)	68.6	71.4	73.3	87.6	75.2	8.5
Lakewood (EK)	76.2	78.1	76.2	73.3	76.0	2.0
Newark (EK)	61.9	63.8	67.6	63.8	64.3	2.4
Lincoln (EK)	77.1	75.2	78.1	81.9	78.1	2.8
Sponsor Scores (N=24):					72.3%	9.1
NFT Scores (N=36):					60.7	9.3
					t = 4.77	
					p < .001	
					f = 4.75	
					p < .01	

Table 5

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--BANK STREET

First Grade

Sites		Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Brattleboro	(EK)	64.4%	74.1%	68.9%	%	69.1%	4.8
Fall River	(EK)	80.7	75.6	71.1	63.0	72.6	7.5
NYC P.S. 243K	(EK)	78.5	74.8	77.0	67.4	74.4	4.9
Philadelphia II	(EK)	77.8	82.6	78.5	77.0	79.0	2.5
Tuskegee	(E1)	80.0	78.5	75.6	76.3	77.6	2.0
Sponsor Scores (N=19):						74.8%	5.5
NFT Scores (N=35):						62.7	6.2
						t = 7.12	
						p < .001	
						f = 2.37	
						p < NS	

Third Grade

Sites		Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Brattleboro	(EK)	75.6%	74.8%	71.1%	%	73.8%	2.4
Fall River	(EK)	61.5	68.9	68.9	71.1	67.6	4.2
NYC P.S. 243K	(EK)	62.2	68.9	77.8	69.6	69.6	6.4
Philadelphia II	(EK)	63.0	65.2	60.0	70.4	64.6	4.4
Tuskegee	(E1)	70.4	81.5	77.0	63.0	73.0	8.1
Sponsor Scores (N=19):						69.5%	6.0
NFT Scores (N=36):						62.4	8.6
						t = 3.20	
						p < .001	
						f = 1.71	
						p < NS	

Mississippi; and Providence, Rhode Island. The greatest difference in implementation scores was found between New York (80.0) and Racine (71.9). In the third grade, the variance is a great within projects as among projects. The difference in New York is particularly great where one third grade classroom has a low score of 57.6 and another has a high score of 81.2. A possible explanation may be that children in University of Oregon's New York third grades have had fewer months in Follow Through and the attrition rate is greater than for other sites. The standard deviation for third grades in St. Louis and Racine is also high. The standard deviation for the first grades at these same sites is considerable less. One possible explanation for the difference among third grade classrooms might be that in the spring of the year when observations are conducted, teachers might be preparing the children for the fourth grade Non-Follow Through classrooms and they might not be adhering so strictly to University of Oregon's stated program. Note, this rationale could also apply to other sponsors as well since there is a slight trend toward lower implementation scores and more deviation among classrooms in the third grade for several other sponsors (see Table 6).

The University of Kansas classrooms differ from the Non-Follow Through classrooms in both first and third grades. The projects assessed were New York City, New York; Philadelphia, Pennsylvania; Portageville, Missouri; Kansas City, Missouri; Louisville, Kentucky. Only one first grade Kansas City classroom's implementation score (64.7) is close to the Non-Follow Through score (62.4). All other University of Kansas classroom scores are in the 70's, 80's, or 90's. The analysis of variance indicates that in the first grade there is a greater difference among site implementation mean scores than there is within sites. Portageville has the highest mean score (92) and Kansas City the lowest (76). Kansas City also has the greatest within-site variance (8.7). The one classroom mentioned above with the 64.7 score seems to account for this variance. In the third grade the greatest variance is found between the two classrooms in New York. The least variation and the highest implementation scores for third grades are found in Kansas City and Louisville (see Table 7).

Overall the High/Scope implementation mean score differs significantly from the Non-Follow Through mean score. The projects assessed are Greenwood, Mississippi; Ft. Walton Beach, Florida; New York City, New York; Greeley, Colorado; Denver, Colorado. Only the classrooms in the New York third grades have implementation scores similar to those in Non-Follow Through. New York also has the lowest first grade implementation scores. The primary difference between New York and the other High/Scope sites is geographic in nature; New York City is the only large eastern urban center included in the High/Scope sample projects. There is little variability within or among site mean scores. In no case is the within-site variance greater than 3.8 (Greenwood first grades) and in Greeley the variance among first grades is only .8. This is remarkable since in Greeley 27 percent of the children speak English as a second language and the attrition rate is high. These figures reflect a migrant, Spanish-speaking population and indicate that the teachers have been able to implement the model in spite of the difficulties which might arise when children speak languages other than the

Table 6

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--UNIVERSITY OF OREGON

Sites	First Grade Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
E. St. Louis (EK)	76.2%	62.5%	76.2%	75.0%	72.5%	6.7
NYC P.S. 137K (EK)	88.7	90.0	91.2		90.0	1.3
Racine (EK)	72.5	72.5	71.2	71.2	71.9	.7
Tupelo (E1)	80.0	86.2	87.5	87.5	85.3	3.8
Providence (EK)	72.5	77.5	72.5	73.7	74.1	2.4
Sponsor Scores (N=19):					78.2%	8.1
NFT Scores (N=35):					61.0	10.7
					t = 6.11	
					p < .001	
					f = 17.61	
					p < .001	

Sites	Third Grade Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
E. St. Louis (EK)	76.5%	62.4%	78.8%	87.1%	76.2%	10.3
NYC P.S. 137K (EK)	68.2	81.2	57.6		69.0	11.8
Racine (EK)	71.8	62.4	84.7	85.9	76.2	11.2
Tupelo (E1)	87.1	80.0	90.6	74.1	82.9	7.3
Providence (EK)	75.0	82.4	78.8	69.4	76.4	5.5
Sponsor Scores (N=19):					76.5%	9.3
NFT Scores (N=36):					60.4	10.5
					t = 5.62	
					p < .001	
					f = .91	
					p < NS	

Table 7

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--UNIVERSITY OF KANSAS

	First Grade					
	Classroom Scores				Site Scores	
Sites	1	2	3	4	$\bar{X}$	S.D.
NYC P.S. 77X (EK)	75.0%	81.3%	%	%	78.1%	4.4
Philadelphia VI (EK)	78.8	90.6	82.4	88.2	85.0	5.4
Portageville (EK)	96.5	91.8	90.6	88.2	91.8	3.5
Kansas City (EK)	82.4	74.1	83.5	64.7	76.2	8.7
Louisville (EK)	85.9	90.6	92.9	85.9	88.8	3.5
Sponsor Scores (N=18):					84.6%	7.9
NFT Scores (N=35):					62.4	8.5
					t = 9.22	
					p < .001	
					f = 5.14	
					p < .01	

	Third Grade					
	Classroom Scores				Site Scores	
Sites	1	2	3	4	$\bar{X}$	S.D.
NYC P.S. 77X (EK)	71.2%	85.0%	%	%	78.1%	9.7
Philadelphia VI (EK)	76.5	82.4	75.3	84.7	79.7	4.5
Portageville (EK)	89.4	74.1	78.8		80.8	7.8
Kansas City (EK)	88.2	88.2	84.7	84.7	86.5	2.0
Louisville (EK)	88.2	91.8	87.1	85.9	88.2	2.5
Sponsor Scores (N=17):					83.3%	6.0
NFT Scores (N=36):					61.3	9.3
					t = 8.89	
					p < .001	
					f = 2.53	
					p < NS	

language used in school. Greenwood, which had a considerably lower per capita income than other High/Scope sites, was also well implemented (see Table 8).

The Education Development Center classroom means differ statistically from the Non-Follow Through classrooms in both the first and third grades. The projects assessed were Burlington, Vermont; Philadelphia, Pennsylvania; Paterson, New Jersey; Rosebud, Texas; Smithfield, North Carolina. The highest implementation score (90) was found in Burlington. Philadelphia has the lowest scores, 65 and 68, and contributes most of the variance for both first and third grades. Not only are Philadelphia implementation scores lower than those of other sites, their within-site variation is greater. Here again, the low implementation scores and variability might be explained by two projonged teacher strikes in Philadelphia. It is possible that when tension is high, teachers may become more structured and adhere less to the theory of the model. It must be noted that all other sites compared to Non-Follow Through have high implementation scores and low within-site variance (see Table 9).

Conclusions

Implementation of the models did not seem to be affected by the percent of children who did not use English as a first language; e.g., the scores in Salt Lake City, Fall River, Greeley, Denver, Fort Worth, and Lakewood were comparatively high.

Average low median family incomes did not seem to be related to successful implementation, e.g., Portageville, Rosebud, Greenwood, and Smithfield. All have low average incomes and high implementation scores.

The size of the city may have had some effect. Large cities accounted for seven of the eight implementation scores below 70. The implementation scores for University of Oregon and High/Scope third grades in New York City were comparatively low and the between-classroom variance in New York was high for University of Oregon, High/Scope, and University of Kansas. In Philadelphia, EDC has low implementation scores for both first and third grades as well as high within-site variance, and Bank Street had low scores for three of the third grades. University of Kansas, however, had high implementation scores and little variation between classrooms in both grade levels in Philadelphia. Perhaps the teacher strikes in Philadelphia made it difficult to implement the more open models represented by Bank Street and EDC since under stress there may be the tendency to become more structured. In Newark all of the first grades scored below the Non-Follow Through mean. The third grades scored close to, but not below, the Non-Follow Through mean. The variation between the classroom scores is very slight. The Newark site has all of the problems of the stereotypical inner-city school; the poverty rate is high, few adults over 25 have a high school education, the community varies in its opinion of the goals for the schools, and the sponsor reported difficulty in providing adequate classroom service during the first several years of sponsorship.

Table 8

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--HIGH/SCOPE

First Grade

Sites		Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Greenwood	(E1)	71.0%	67.6%	70.3%	75.6%	71.4%	3.8
Ft. Walton Beach	(E1)	77.9	73.8	79.3	75.3	76.7	2.4
NYC P.S. 92M	(EK)	66.2	71.7	71.0		69.7	3.0
Greeley	(EK)	82.8	81.4	82.8		82.3	.8
Denver	(EK)	86.9	80.7	80.7	82.8	82.8	2.9
Sponsor Scores (N=18):						76.6%	6.0
NFT Scores (N=35):						63.7	5.8
						t = 7.58	
						p > .C01	
						f = 15.59	
						p > .001	

Third Grade

Sites		Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Greenwood	(E1)	70.3%	73.1%	75.2%	74.5%	73.3%	2.1
Ft. Walton Beach	(E1)	83.4	83.4	80.7	78.6	81.6	2.3
NYC P.S. 92M	(EK)	66.9	62.1	64.8	64.8	64.7	2.0
Greeley	(EK)	80.0	80.0	86.2		82.1	3.6
Denver	(EK)	73.1	71.7	78.6	76.6	75.0	3.2
Sponsor Scores (N=19):						75.0%	6.9
NFT Scores (N=36):						63.5	6.8
						t = 5.93	
						p > .001	
						f = 27.34	
						p > .001	

Table 9

TOTAL IMPLEMENTATION SCORES FOR CLASSROOMS BY SITE--EMC

Sites	First Grade Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Burlington (EK)	91.0%	91.0%	86.0%	93.0%	90.2%	3.0
Philadelphia IV (EK)	72.0	80.0	58.0	48.0	64.5	14.3
Paterson (EK)	70.0	79.0	79.0	73.0	75.2	4.5
Rosebud (EK)	72.0	68.0	71.0		70.3	2.1
Smithfield (E1)	85.0	76.0	86.0	83.0	82.5	4.5
Sponsor Scores (N=19):					76.9%	11.5
NFT Scores (N=35):					61.2	9.6
					t = 5.35	
					p > .001	
					f = 7.26	
					p > .01	

Sites	Third Grade Classroom Scores 1	2	3	4	Site Scores $\bar{X}$	S.D.
Burlington (EK)	85.5%	81.8%	79.1%	75.5%	80.5%	4.2
Philadelphia IV (EK)	75.5	64.5	73.6	59.1	68.2	7.7
Paterson (EK)	67.3	69.1	77.3	73.6	71.8	4.5
Rosebud (EK)	85.5	80.0	79.1		81.5	3.4
Smithfield (E1)	76.4	79.1			77.7	1.9
Sponsor Scores (N=17):					75.4%	7.1
NFT Scores (N=36):					60.7	10.6
					t = 5.18	
					p > .001	
					f = 4.54	
					p > .05	

In general, small sites such as Smithfield, Greeley, Portageville, Tupelo, Tuskegee, and Lebanon were implemented more consistently at both grade levels and they had higher implementation scores in both grades. This seems to indicate that the educational change represented by the Follow Through sponsors may be instituted more successfully in the smaller cities and towns than in the large inner-city schools. There are exceptions, of course; University of Kansas was successful in both grade levels in Philadelphia.

Overall it is possible to conclude that with minor exceptions the seven Follow Through sponsored models considered in this study have been implemented as planned.

Sponsor Differences

In an effort to see whether sponsors do differ from one another in classroom instructional processes, a discriminant function analysis was conducted. The discriminant functions on which sponsors differed were dominated by one or two very specific classroom process variables, e.g., "Large group with an adult in math" or "Adult reinforcement with a token in an academic subject." Three groups of classrooms were usually found in these analyses. Those classrooms using the University of Kansas model formed a cluster around the variables describing the use of tokens, those using the University of Oregon model formed another cluster around the variables using an adult with a large group in math, and classrooms of the remaining five sponsors formed several different clusters.

A separate analysis was conducted to see whether the five sponsors could be distinguished from each other if the data from University of Kansas and University of Oregon were not used. In this analysis, the University of Arizona was distinguished from the remaining four sponsors on the basis of "Child's extended response to questions" and "Adult communication or attention focus, small group." High/Scope was differentiated from the other four on the basis of variables indicating a high level of verbal interaction between adults and children and, as would be expected from their Piagetian model, "Children self-instruction, objects." The third discriminant function distinguished Far West Laboratory from EDC and Bank Street in that Far West had more "all adult praise to children" and less "academic instruction."

An analysis was also made to see if classrooms could be classified by sponsor. Based on classroom observation data, out of a total of 524 classifications, 410 were correct. University of Kansas and University of Oregon classrooms were rarely misclassified as belonging to another sponsor. The classrooms of the remaining five sponsors are occasionally confused with each other but only rarely with the University of Oregon or University of Kansas models. In the great majority of cases, however, classrooms affiliated with a particular sponsor were correctly identified with that sponsor and we conclude that, for the most part, sponsors can be distinguished by the observation variables used in this analysis.

In order to learn more about the type of processes used in the Non-Follow Through classrooms, their scores on the critical variables were also assessed to see how they would be classified in the sponsors' groups. Few Non-Follow Through classrooms were classified as University of Kansas classrooms in either grade level. They were most often classified as EDC in the first grade and as University of Oregon in the third grade, on the grouping and activity variables. On the interaction variables, the Non-Follow Through classrooms were distributed rather evenly across Far West, University of Arizona, Bank Street, University of Oregon, High/Scope, and EDC.

Classroom Instructional Processes and Child Outcomes

The study of implementation would be of little importance if we did not believe that differing educational theory and practices affect children differently.

Like educators in general, Follow Through sponsors feel that the development of basic skills in reading and computing is important, but that it is also desirable for children to develop such attributes as task persistence, attending ability, cooperation, inquiry behavior, and independence. While these attributes appear to be illusive, we have been able to operationally define and systematically observe some of these behaviors.

Sample

Of the classrooms observed, a total of only 105 first grades and 58 third grades met the criterion of having both baseline and Spring 1973 test scores, and only these classrooms were included in the study of classroom processes and child outcomes. The smaller number of third grade classrooms reflects the attrition of children with baseline data. Table 10 indicates the number of classrooms per sponsor included in the study. All sponsors' classrooms and Non-Follow Through classrooms which were both observed and had sufficient baseline data were merged in the study. Thus, the study is one which examines classroom processes regardless of model and relates the processes to child outcomes. This procedure provides a wide range of classroom processes to be examined. The unit of analysis in this study was the classroom.

Statistical Procedures

To examine relationships between observed classroom practices and child outcomes, partial correlations were computed, holding constant the baseline WRAT scores. Table 11 describes the process variables used in these computations. Stepwise regressions were computed using selected observed process variables and all child outcomes; the WRAT entered the regression equation first.

Table 10

NUMBER OF CLASSROOMS AND SITES INCLUDED IN THE PARTIAL CORRELATIONS AND REGRESSION ANALYSES BY SPONSOR AND GRADE LEVEL

Sponsor	First Grade		Third Grade	
	Number of Classrooms	Number of Sites Represented	Number of Classrooms	Number of Sites Represented
Far West Labs	12	3	14	4
University of Arizona	14	4	2	1
Bank Street College	11	4	7	2
University of Oregon	5	2	4	1
University of Kansas	17	5	12	4
High/Scope	13	4	0	0
EDC	12	3	6	2
Non-Follow Through	24	-	13	-
Total Number of Classrooms	108		58	

Table 11

PARTIAL CORRELATION ANALYSES

	Number of Process Variables	Number of Classrooms
Child Behaviors		
First Grade	28*	105
Third Grade	28	58
Days Absent		
First Grade	340	108
Third Grade	340	58
Raven's--Third Grade	340	58
Coopersmith--Third Grade	340	58
IAR--Third Grade	340	58
MAT		
First Grade	340	108
Third Grade	340	58

*The 28 variables are a subset of the 340 variables used in the other analysis.

Reading Achievement Results

Out of a possible 340 correlations between reading achievement and classroom processes, 118 were significantly related at the .05 level. Of these, the most strongly correlated variables suggest that the length of the school day and the average time a child spent engaged in a reading activity were related to higher reading scores in both first grade and third grade. When the school day is longer, the children have more opportunity to engage in reading. The length of the school day for the classrooms in the evaluation varied among schools by as much as two hours. Higher reading scores were also found in classrooms where there was more reading or discussions of reading between adults and children. Thus, opportunity and exposure to reading had an important relationship to good performance on tests.

Higher reading scores were obtained in classrooms using systematic instructional patterns where the teacher provides information and asks a question about the information. The child responds and the teacher immediately lets the child know whether the response is right or wrong. If he is wrong, the child is guided to the correct answer. If he is correct, he receives praise, a token, or some form of acknowledgment. These preliminary findings suggest this type of positive reinforcement contributed to higher reading test scores in both first and third grades.

Small groups were most effective for teaching first grade reading, while large group instruction worked well in the third grade. In classrooms where children worked by themselves and were task persistent (maintained their attention on their studies without teacher guidance), they also achieved higher reading scores. In classrooms where textbooks and programmed workbooks were used most often, the reading scores were higher. Also, in classes where a greater-than-average amount of time was spent on social studies, the reading scores were higher. Obviously, reading skills are used in social studies projects, but it is of interest to note that experience in social studies was related to reading scores.

It is noteworthy that the University of Oregon and the University of Kansas, both of which are models that use the classroom procedures described here, showed greater gains in first grade reading than the other five sponsors and greater gains than Non-Follow Through classes.

Math Achievement Results

Out of a possible 340 correlations between math achievement and classroom processes, 108 were significantly related at the .05 level. Of these, the most strongly correlated variables suggest that, as in reading, the length of the school day and the average length of time each child spent in math activities were related to higher math scores in both first and

third grades. Thus, the opportunity a child had to engage in math, either in formal instruction or in less formal exploratory activities (e.g., working with, or just "messing with," weights and measuring tools) contributed to higher scores in math. Also, in classrooms where adults and children more often discussed or talked about mathematical problems and concepts, the test scores in math were higher. The value (in terms of math scores at the end of the third grade) of spending large amounts of class time on math was especially marked for the children whose numerical ability was weak when they entered school.

The effect of praise on achievement in math in first grade was variable: in classrooms where children had relatively low entering ability, the children profited more from a high rate of praise than they did in classrooms where the students had higher entering ability. This type of information could be useful in planning educational programs to enhance the learning of children with differing abilities at different age levels.

As in reading, children had higher math scores in classrooms where teachers used systematic instructional patterns; that is, the teacher provides information and asks a question about the information. The child responds and the teacher immediately lets the child know whether the response is right or wrong. If he is wrong, the child is guided to the correct answer. If he is correct, he receives praise, a token, or some other form of acknowledgment. This positive reinforcement contributed to higher scores on math tests in both grade levels.

In classrooms where textbooks and programmed workbooks were used frequently, the test scores on math were especially high. In addition, the use of instructional materials such as programmed materials, Cuisenaire rods, or Montessori materials contributed to higher math scores.

In first grade classrooms where children were taught in small groups, the math scores were higher. In third grade, large group instruction contributed to higher scores. When children could work by themselves some of the time and could persist at a task, they were also more likely to have higher scores in math achievement.

University of Kansas, which used the classroom procedures described here as contributing to higher math scores, had higher scores in first grade math than the other six sponsors and Non-Follow Through classes. University of Oregon, which also used these instructional processes in their classrooms, had higher scores in the third grade math than the other six sponsors and Non-Follow Through classes. These findings strongly suggest that classroom procedures used in University of Kansas and University of Oregon classrooms contributed to child achievement in math.

Raven's Problem Solving Test Results

Out of a possible 340 correlations between the Raven's Problem Solving Test and classroom processes, 114 were significantly related at the

.05 level. Of these, the most strongly correlated variables suggest that high scores on Raven's Coloured Progressive Matrices (a test of nonverbal perceptual problem solving) tended to be earned by children in the more flexible classrooms where a wide variety of materials are used, many different activities occur, and children are allowed to select their own groups and seating part of the time. In these more flexible classrooms, children have more opportunities to manipulate materials and discover the relationships between items to see how things fit together. In these classrooms, adults interact with children on a one-to-one basis, more open-ended questions are asked, and children show more verbal initiative. Far West Laboratory, University of Arizona, Bank Street College, High/Scope Foundation, and Educational Development Center use these processes, and the classrooms in these models had higher scores on the Raven's than did the classrooms in the University of Kansas and University of Oregon models.

Responsibility Scale Results

Out of a possible 340 correlations between the Intellectual Achievement Responsibility Scale and classroom processes, 106 were significantly related at the .05 level. Of these, the most strongly correlated variables suggest that children in the more open classrooms earned higher scores on the Intellectual Achievement Responsibility Success Scale. Our results indicate that children from the more flexible classrooms took responsibility for their own success but not for their failure. Children from the more highly structured classrooms took responsibility for their own failure but attributed their success to their teacher's competence or other forces outside themselves. Only the classrooms of Educational Development Center had scores indicating that the children took responsibility for both their success and failure.

Days Absent Results

The absence rate is important for several reasons; e.g., many school budgets are determined by the average daily attendance. Also, days absent can be used as an indicator of attitude toward school. It is well known to parents and teachers that if a child enjoys school, he may attend even if he does not feel very well. If he does not like school, he is more likely to stay home whenever he feels any discomfort.

Out of a possible 340 correlations between days absent and classroom processes, 102 were significantly related at the .05 level. Of these, the most strongly correlated variables suggest that in both first and third grade classrooms, children are absent less frequently in open classrooms--that is, in classrooms where there is a high rate of child independence, child questioning, adults responding, individualized instruction, and open-ended questioning. Also, in classrooms where children and adults smiled and laughed more often, the children were absent less often.

Children in both first and third grade were absent more frequently from classrooms where they worked in large groups more often and where adults

used direct questions in academic work and frequent corrective feedback. Findings for the third grade indicate that in classrooms where children were punished they also were absent more often. In addition, classrooms with a higher rate of negative, harsh, or demeaning statements on the part of teachers and students showed a higher absence rate.

The findings in this report of absence rate indicate that at the first grade level, children in classrooms of sponsors who used more highly structured environments, materials, and interactions also had a higher absence rate. Classrooms of three sponsors, Far West Laboratory, University of Arizona, and High/Scope Foundation, models which used a wide variety of activities and materials, had children who had lower absence rates than children in classrooms of other sponsors and Non-Follow Through classrooms. As might be expected, the absence rate for all sponsors and Non-Follow Through diminished from first grade to third grade.

Child Behaviors Results

Table 13 presents the results of the partial correlations for child independence, task persistence, cooperation, and question asking.

Independence--In our study, independence is defined as a child or children engaged in a task without an adult. This type of independent behavior is more likely to be found in classrooms where teachers allow children to select their own seating and groups part of the time, where a wide variety of activities is available, and where an assortment of audiovisual and exploratory materials is available. The adults provide individual attention and make friendly comments to the children.

Our investigations indicate that children in the classrooms of Educational Development Center and Far West Laboratory showed more independence than did the children in Non-Follow Through and the other five sponsors' classrooms.

Task Persistence--For this study, task persistence is defined as a child engaged in self-instruction over a few minutes or more. If the child becomes engaged in a conversation with someone else during the task, the observer no longer codes task persistence. The highest positive relationships indicate that task persistence occurred most often when textbooks and workbooks were used in the classroom. Where adults instructed one child at a time, the children were also likely to be more task persistent. This may be because young children often have difficulty understanding group instructions. However, in settings where adults work with children on a one-to-one basis, children can have a question answered or directions clarified and then go ahead independently with the task at hand.

University of Arizona and University of Kansas had higher scores on task persistence than do the other five models and Non-Follow Through.

Table 13

PARTIAL CORRELATIONS OF INSTRUCTIONAL VARIABLES
AND CHILD BEHAVIORS
(Fall 1971 WRAT Partialed Out)

Instructional Variables	Correlations			
	Independence	Task Persistence	Cooperation	Child Questions
Child/Adult Ratio	.23*	.09	.02	-.15
Children Select Groups and Seats Part of the Time	.36***	-.22*	.19*	.03
Instructional Materials Used	-.01	.11	.09	-.07
Audiovisual Equipment Used	.13	-.25**	.15	-.12
General Equipment and Materials	.22*	-.08	.09	.005
Total Resource Materials Used	.13	-.23*	.18	.03
Wide Variety of Activities Occur Concurrently	.22*	-.12	.15	.09
Wide Variety of Activities Occur During the Day	.43***	-.36***	.32**	.14
An Adult with One Child	.57***	-.16	.08	.14
Use of TV	-.03	-.10	-.11	-.03
Audiovisual Equipment Used in Academic Subjects	.24**	-.25**	-.01	-.04
Exploratory Materials Used in Academic Subjects	.34***	-.22*	.27**	-.11
Math or Science Equipment Used in Academic Subjects	-.18	.17	-.18	.11
Textbook and Workbooks Used in Academic Subjects	-.33***	.31**	-.49***	-.04
Puzzles and Games Used in Academic Subjects	.16	-.07	.09	-.07
Adults Asking Children Questions	-.17	.03	-.17	-.04
Adult Instructs an Individual Child	-.09	.23*	-.17	.22*
Adult Comments to Children	.22*	-.12	-.13	.36***
Adult Task Related Comments to Children	.12	-.24*	.39***	-.16
Adult Acknowledges Children	-.16	.15	-.11	.04
Adult Praises Children	-.60***	.20*	-.21*	.02
Adult Speaks to One Child	-.01	.13	-.06	.38***
Adult Speaks to Two Children	.29**	-.13	.28**	-.03
Adult Speaks to a Small Group	-.15	.19*	.01	-.32***
Adult Asks Direct Question about Subject Matter	-.41***	.07	-.28**	.03
Adults Ask Open-Ended Thought-Provoking Questions	.16	-.12	.13	-.07

* $p < .05$
** $p < .01$
*** $p < .001$

Number of classrooms used in the correlation computations = 105 first grades.

Cooperation--For this study, cooperation is defined as two or more children working together on a joint task. This kind of cooperation is more likely to be found in classrooms where a wide variety of activities occur throughout the day, where exploratory materials are available, and where children can choose their own groupings. If the adults interact with two children, asking questions and making comments about the task, the children seem to be encouraged to join each other in cooperative tasks.

The children in the Bank Street College, High/Scope Foundation, and Educational Development Center programs more often joined each other in a cooperative task than did children in the other four models and Non-Follow Through children.

Question Asking--Educators have long recognized the value of a child's asking questions as a primary means to gain information. Previous research indicates that question asking is positively related to test scores.[3] In our study, we found that first grade children asked more questions where there was a one-to-one relationship of adult with child in classrooms, where adults responded to children's questions, and where adults made general conversational comments to children.

Children in classrooms using Far West Laboratory, Bank Street College, University of Kansas, High/Scope Foundation, and Educational Development Center programs ask questions more often than do children in the Non-Follow Through classrooms.

Child Outcome Scores Explained by Entering Ability[4] and Classroom Processes

Whether or not classroom procedures affect the growth and development of children has been seriously questioned by other research (Coleman, Jencks, Herrnstein, Moynihan, and Mosteller). Their research has indicated that a child's entering aptitude is of primary importance and, in fact, governs what the child will achieve in school. The study reported here, however, found that observed classroom procedures contributed as much to the explanation of test score differences as did the initial ability of children. Table 14 presents findings from a stepwise regression where the WRAT score was entered into the regression first. The third and seventh columns report that part of the variance explained uniquely by the process variable.

In both first and third grades, child behavioral outcomes were only slightly explained by entering aptitude. As might be expected, these behaviors were much more related to classroom processes.

Very little of the absence rate was explained by entering ability, in either first or third grade. Approximately 60 percent of the variance was explained by the instructional procedures used in the classroom, suggesting that what occurs in classrooms is related to whether or not the child stays away from school.

Table 14

SUMMARY STATISTICS FOR THE STEPWISE REGRESSION ANALYSES

Outcome Variable	First Grade (N=105) R^2_{WRAT}	First Grade (N=105) $R^2_{WRAT, Process\ Variables}$	First Grade (N=105) Process Variables (Unique)	First Grade (N=105) Number of Process Variables (included in regression)*	Third Grade (N=58) R^2_{WRAT}	Third Grade (N=58) $R^2_{WRAT\ Process\ Variables}$	Third Grade (N=58) Process Variables (Unique)	Third Grade (N=58) Number of Process Variables (included in regression)*
Behavioral Outcome Variables								
Child Questions	.00	.28	.28	2	.00	.29	.29	3
Self-Esteem	.00	.48	.48	5	.01	.42	.41	6
Child Independence	.00	.67	.67	5	.01	.41	.40	3
Task Persistence	.00	.44	.44	9	.06	.61	.55	6
Cooperation	.00	.32	.32	2	.10	.61	.51	5
Verbal Initiative	.00	.22	.22	3	.00	.38	.38	2
Days Absent	.01	.67	.66	14	.07	.69	.62	6
Test Outcome Variables								
MAT Math	.32	.74	.42	10	.17	.81	.64	8
MAT Reading	.50	.73	.23	8	.42	.79	.37	7
Raven's	--	--	--	--	.41	.86	.45	9
IAR-Success	--	--	--	--	.18	.57	.39	4
IAR-Failure	--	--	--	--	.04	.83	.79	11

*This column contains the number of process variables that entered the stepwise regression with an "F-to-enter" that was significant at .05.

The achievement of a child in math at the end of first grade can be attributed in part to his ability as it was measured when he entered school, but even more so by the instructional practices used by his teachers. In first grade, entering ability accounts for approximately 40 percent of the achievement (Table 14). By the third grade, less of the achievement can be attributed to entering school ability and more to classroom practices. Table 15 lists those process variables which entered the stepwise regression.

In first grade we found that a variable which describes a stimulus/response/feedback (S/R/F) sequence of interaction entered the regression equation after the WRAT and explains 13 percent of the variance of the math scores. Eight of the 10 variables which entered the equation are related to this S/R/F sequence.

In third grade, 25 percent of the test score variance is explained by the process variables which describe adults asking children questions about academic subject matter. The WRAT only explains 17 percent of the variance.

Approximately 50 percent of first grade reading achievement can be attributed to the entering ability of the children. The instructional procedures used by teachers account for approximately 25 percent of the reading achievement. The variables which entered the equation are listed on Table 16.

In the first grade, the total number of verbal interactions which were related to reading accounted for 12 percent of the variance in first grade reading scores. The other variables which entered the equation were primarily related to average amount of time spent in reading and stimulus/response/feedback variables.

In third grade, reading success can be attributed about equally to the instructional procedures used by teachers and the entering ability of the children.

Table 16 displays data that indicates that an adult working with a large group of children accounts for 16 percent of the third grade reading score variance. Total academic verbal interaction accounts for less of the variance (4 percent) in third grade than in first grade. This may be explained by the fact that third grade children may not need as much interaction with adults about reading and work more on their own.

One of the most important findings centers around the Raven's test of nonverbal reasoning or perceptive problem solving (considered to be a culture-fair test of fluid intelligence). The abilities required to function well on this test have not been considered to be influenced by environment. This study found that ability to perform well on the Raven's test was related to the classroom environment and strongly suggests that children who, for a period of three years, have been in classrooms that use a wide variety of activities and provide a wide variety of manipulative materials have leanred to see the relationship between parts and wholes. At any rate, they learn to see spatial relationships similar to those tested on the Raven's.

Table 15

STATISTICS FROM THE STEPWISE REGRESSION OF THE BASELINE WRAT AND SELECTED COI VARIABLES ON THE OUTCOME SCORES

	Variable Name	Multiple R	R^2	RSQ Change	F-to-Enter
Dependent Variable: WRAT Math					
First Grade (N = 105)					
1	F71 WRAT	.59	.32	.35	56.10
2	Adult feedback to children responding to adult academic command or question	.70	.49	.13	26.32
3	Child extended response, nonacademic	.74	.55	.07	14.57
4	Adult positive corrective feedback, academic	.77	.59	.04	10.79
5	Adult communication or attention focus, group	.79	.63	.03	8.31
6	All adult reinforcement with tokens	.81	.65	.03	8.20
7	Child extended responses, academic	.83	.68	.03	8.07
8	Adult communication, request, and direct question to individual child, academic	.84	.70	.02	6.39
9	Social interaction	.85	.71	.01	4.54
10	Child responses, academic	.85	.73	.01	4.18
11	Adult neutral corrective feedback, academic	.86	.74	.01	4.24
Third Grade (N = 57)					
1	Classroom mean F69 WRAT	.41	.17	.17	11.10
2	Adult academic command, request, and direct question to children	.64	.41	.25	22.48
3	Large group with any adults	.74	.55	.14	16.39
4	Child self instructs, academic, with objects	.79	.62	.06	8.69
5	Child presents information to a group	.82	.67	.05	8.20
6	Numbers, math, arithmetic	.85	.72	.05	8.87
7	Child selection	.87	.76	.04	8.48
8	Social interaction	.89	.79	.03	6.32
9	Child's extended response to question	.90	.81	.02	4.62

Table 16

STATISTICS FROM THE STEPWISE REGRESSION OF THE BASELINE WRAT AND SELECTED COI VARIABLES ON THE OUTCOME SCORES

	Variable Name	Multiple R	R^2	RSQ Change	F-to-Enter
Dependent Variable: MAT Reading					
First Grade (N = 105)					
1	F70 WRAT	.71	.50	.50	103.40
2	Total academic verbal interaction	.79	.63	.12	33.89
3	Total weight in reading groups	.81	.65	.02	6.47
4	Individual child response to adult academic command, request, or direct questions	.81	.66	.01	3.57
5	All adult reinforcement with tokens	.83	.68	.02	6.13
6	Child extended response, academic	.83	.70	.02	5.01
7	Adult neutral corrective feedback, academic	.84	.71	.01	4.62
8	Teacher with large group	.85	.72	.01	3.72
9	Adult with large group	.86	.73	.01	4.42
Third Grade (N = 57)					
1	F70 WRAT	.65	.42	.42	40.01
2	Large group with any adults	.76	.58	.16	20.75
3	Child/teacher and aide ratio	.79	.62	.04	5.25
4	Total academic verbal interaction	.81	.66	.04	6.68
5	Total class duration	.84	.70	.04	6.48
6	Small group with any adults	.87	.75	.05	10.07
7	Child selection	.88	.77	.02	4.45
8	Large group with any adults	.89	.79	.02	4.08

Conclusions

A study of the instructional procedures used in classrooms and the achievement of children indicates that time spent in reading and math activities and a high rate of drill, practice, and praise contribute to higher reading and math scores. Children taught by these methods tend to accept responsibility for their failures but not for their successes. Lower absence rates, higher scores on a nonverbal problem solving test of reasoning can be attributed in part to more open and flexible instructional approaches in which children are provided a wide variety of activities and materials and where children engage independently in activities and select their own groups part of the time.

Classroom instructional processes predicted as much or more of the outcome score variances than did the entering school test scores of children. Based upon these findings, we conclude that what occurs within a classroom does contribute to achievement in basic skills, good attendance, and desired child behaviors.

Footnotes

1. Stallings, Jane, and David Kaskowitz, Follow Through Classroom Observation Evaluation 1972-1973, Menlo Park, CA: Stanford Research Institute, 1974.

2. Stallings, Jane A., and Phillip A. Giesen, A Study of Reliability in Observational Data, Menlo Park, CA: Stanford Research Institute, 1974.

3. Stallings, Baker, and Steinmetz (1972) and Stallings (1973) report that an increased frequency of children asking questions is related to higher scores on achievement tests and attitudinal tests.

4. Measured by the Wide Range Achievement Test administered when the child entered school.